AN EXAMINATION OF MODELLING APPROACHES FOR PREDICTING THE DISPERSION OF AEROSOLS FROM OBSCURANT GRENADES

W.S. Andrews and **S.D**. Roney
Department of Chemistry and Chemical Engineering
Royal Military College of Canada
Kingston, Ontario, Canada K7K 5L0

Gilles Roy
Defence Research EstablishmentValcartier
Val-Belair, Quebec, Canada G3J 1X5

An artificial neural network model (ANN) has been developed to predict the aerosol concentration distribution resulting from the volley firing of screening grenades for armoured vehicles. The data were collected during trial firings using the Defence Research Establishment Valcartier (DREV) Laser Cloud Mapper (LCM), a scanning LIDAR device operating at 1.06 μm. The ANN model was able to predict volumetric extinction coefficients, and thus aerosol concentrations, to within an average error of 1.78 for a validation set of 1103 data points. A basic Gaussian plume model was also used to generate predictions over the same validation set, but the average error was 14.5, almost an order of magnitude worse than the ANN predictions. An examination of the extinction coefficients along a line-of-sight normal to the aerosol cloud revealed a bimodal distribution, which was also predicted by the ANN model. The Gaussian model predicted a monomodal distribution.

As the data set used for training, testing and validation was limited to the firing of two volleys of different composite grenades (phosphorus and metal flake), the ANN model can not be considered adequate for general application, but rather serves as a proof of principle that neural networks are capable of predicting the spatial and temporal distribution of aerosol particles better than conventional Gaussian models. A program of further data collection at DREV and subsequent analysis using ANNs is planned for the future.

Introduction

For some years, Defence Research Establishment Valcartier has used a scanning light detection and ranging (LIDAR), called a Laser Cloud Mapper (LCM)[1] to measure the attenuation of laser pulses caused by various aerosols under consideration for military use.[2,3,4] The backscatter data have been converted to volume extinction coefficient maps and give an indication of the concentration distribution of the obscurant aerosols. Although these data, collected over several years and for a wide variety of aerosols, were mainly used to assess obscuration, they also present information on the dispersion of aerosols from a point source which emits over a relatively short duration (<60 see). It is sig-

nificant to note that obscuring grenades, which form a large portion of the data set, are neither instantaneous (puff) nor continuous (plume), but rather somewhat transient, with a short burning/emission time.

The classical prediction models are the Gaussian puff for instantaneous emissions and the Gaussian plume for continuous emissions. This paper examines the adequacy of Gaussian models to predict the aerosol dispersion measured by the LCM and to compare these to a model developed using artificial neural networks (ANNs), an aspect of artificial intelligence and a relatively new approach to modelling.

Gaussian Modelling

The two Gaussian models, the plume and the puff, are based on diffusion theory, with the concentration decreasing exponentially from the centreline or centre.[5] The more widely used, and thus the better characterized of the two is the plume model, which has a variety of applications for aerosol and vapour dispersion. The general Gaussian plume equation describing aerosol concentration, C (g/m^3), from a continuous source, is

$$C = \frac{Q}{2\pi\sigma_y\sigma_z u}\left(e^{-\left(y^2/2\sigma_y^2\right)}\right) \times$$
$$\left[e^{-(z+h)^2/2\sigma_z^2} + e^{-(z-h)^2/2\sigma_z^2}\right] \quad (1)$$

where Q is the source/production rate (g/s), σ_y and σ_z are standard deviations in azimuth and elevation respectively (m), y and z are respective azimuthal and vertical distances from the centreline (m), h is the distance of the plume centreline above ground (m) and u is the wind speed (m/s) taken as being along the x axis.

A Gaussian distribution, by definition, has two fitting parameters, a mean (in this case the centreline concentration provided by the first term on the RHS of the Eq. 1) and a standard deviation (here contained in the σ_y and σ_z terms). It is noteworthy that for the aerosol dispersion application these standard deviations are determined from the appropriate Pasquill stability categories which in turn are functions of distance from the source (x), cloud cover, time of day, etc. The basic assumptions inherent in the Gaussian plume model are constant emission, conservation of mass, (no ground deposition), steady wind and that the results are time averaged. Such parameters as ambient temperature, relative humidity, surface roughness as well as ground deposition any chemical kinetics or aerosol/atmosphere reactions are not explicitly or even implicitly included in the basic model, although additional terms can be added. Two further limitations are that the Gaussian plume model predicts rather poorly less than 100 m from the source and that the quality of predictions varies inversely with distance from the centreline.[7]

Although some of the limitations mentioned above could be addressed by physically-based or semi-empirical fittings (notwithstanding that the σ_y and σ_z values are themselves semi-empirical in nature), it was decided to use the general Gaussian plume model as the basis of comparison for ANN models.

The instantaneous Gaussian puff model, which intuitively should provide a better prediction of aerosol dispersion from screening grenades than the con-

tinuous plume model, was found to perform not nearly as well. This is probably due to the fact that the wider use of the plume model has led to much better characterizations of the standard deviations (σ_y and σ_z) values used for predictions.[8]

Artificial Neural Network Models

Artificial neural networks are considered a branch of artificial intelligence and found their origins in psychologists attempts at modelling the brain's functions. The seminal work on the back propagation paradigm, the most widely used of ANN models, was only published in 1986.[9]

Basically, an artificial neural network takes raw input data, maps it into an input space, and then introduces it into the network. The network itself is an interconnection of nodes or processing elements (also sometimes called artificial neurons) which are arranged in layers (Fig. 1). The input layer contains a node for each variable, with each complete set of variables describing a specific input vector. For example, each measured value of volumetric extinction coefficient has a corresponding vector of measured experimental variables and parameters. In addition to the input variables, the input layer also contains a bias node which is permanently assigned a value of 1 and is analogous to an electrical ground. Its presence was found to aid convergence during training.

Each node in the input layer is connected to each node in the next layer, which is usually called the hidden layer. The scaled (mapped) input values each have a weight applied to them before the values are introduced into the hidden layer. At each hidden node, the weighted inputs are summed and then mapped back into the range -1 to +1 by the application of a hyperbolic tangent transfer function before being passed to the output layer. Conventionally, there are one or two hidden layers, with the interconnections between nodes in the two hidden layers being analogous to those between the input and first hidden layer.
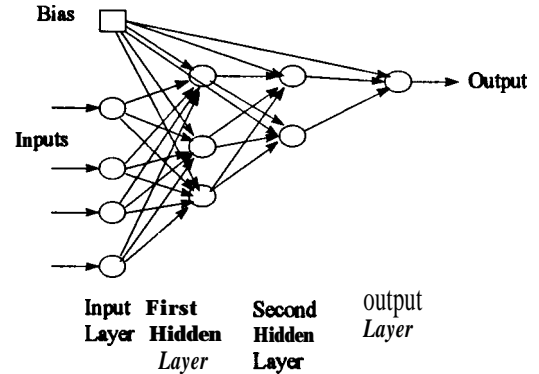

Fig. 1. Typical architecture of a back propagation neural network.

Again, each hidden node in the final hidden layer, as well as the bias node, is connected to each output node, with the process of applying weights to each connection, the weighted inputs being summed and then having a transfer function applied being repeated. The output of the output node(s) must then be mapped back into real-world value(s).

To initial scaling of input variables, is effected by:

$$ x_i = 2V_i \left(\frac{M_i + m_i}{M_i - m_i} \right), \qquad (2) $$

where V_i is the unscaled i^{th} input value, M_i is the maximum value of the range of the ith input, m_i is the minimum value of the range of the ith input and x_i is the scaled i^{th} input value.

The scaled inputs (which can also be considered as the outputs of the input layer) have the appropriate connecting weights, w_{ji} applied. These weighted outputs become the inputs for the nodes of the first hidden layer, where they are summed and then transformed using a transfer function, as outlined by the following equations:

$$I_j = \sum_{i=0}^{n} w_{ji} x_i \; , \qquad (3)$$

where w_{ji} is the connection weight between the j* hidden node and the i* input node, and I_j is the sum of the weighted inputs to the j* hidden node, and

$$y_j = \tanh(I_j) = \frac{e^{I_j} - e^{-I_j}}{e^{I_j} + e^{-I_j}}, \qquad (4)$$

where y_j is the output of the jth hidden node.

This process is repeated for the second hidden layer and the output layer. The output value from the output layer must then be mapped back to provide a real value for the logarithm of the volumetric extinction coefficient. In fact, this process is essentially the reverse of the initial scaling described in Eq.2, although it is simplified by only having to consider the mapping of one variable. This is accomplished by

$$C = \frac{(M-m)z + (Rm - rM)}{R - r} \qquad (5)$$

where C is the predicted value of the logarithm of the volumetric extinction coefficient corresponding to the specific input vector defined by $X_1, X_2 .. .x_g$, M and m are the real (measured) maximum and minimum of the output variable (log σ), z is the network output value, whose range falls between the maximum and minimum values R and r (for the tanh function, these values are taken as 0.8 and -0.8 respectively).

At the start of the training process, the connecting weights throughout the network are given random values. When the output value is generated, it is compared to the corresponding actual measured value. The difference between these two values is considered the global error and is propagated backwards through the network. An iterative process is carried out of adjusting the connecting weights to minimize the global error. A gradient descent optimization is used to adjust each of the connecting weights locally. Once the global error has been minimized over the whole training set, the weights can be fixed and the network can be used to make blind predictions.

A trained ANN lends itself particularly well to incorporation into a spreadsheet which can be driven by a macro. It can also be incorporated into computer codes to provide real-time determination of the extinction coefficients and thus aerosol concentrations or transmittances, as required.

Data Collection

Some of the trials involving the LCM include Smoke Week IX[10], SOCMET Winter[11] and Summer[12]. The data used in this study were for two experimental screening grenades, each composed of a mix of red phosphorus (RP) and metal flake. The LIDAR was situated some 400 m from the cloud centreline so that the field of view was roughly perpendicular to the longitudinal

axis of the cloud. The grenades were 66 mm in diameter and designed for armoured fighting vehicle launchers. In each case, 12 grenades were launched together, with their impact over an arc of about 70 m. The red phosphorus was actually ignited in the air during flight through the launch arc. The metal flake was also dispersed while the grenades were in the air (at an elevation of some 10 m). As the grenades continued to burn and emit smoke after the rubberized RP particles had reached earth, the source was considered to be a ground release (h=O in Eq. 1). Even though this was not strictly true, it was not felt to contribute significantly to any modelling errors, as the data for the modelling were not collected until a coherent cloud had formed.

The scanning LIDAR is based on a 1.06 μm NdYAG laser. Although the settings are adjustable, those used for the trial from which the data were collected were such that, for one complete scan or cycle, the LCM traversed the 90° field of view six times, making 66 shots per traverse (Fig.2). The vertical field of view was 2°. The LCM measures the backscatter from the laser pulses, with a sampling frequency of 100 MHz. The resolution available was 1.5 m along the shot axis, 9.5 m at 400 m in azimuth and 2.8 m at 400 m in elevation. A typical scan would produce some 198 000 data points (backscatter values).

Electromagnetic attenuation by atmospheric or artificial aerosols can be characterized by the Beer-Lambert Law:

$$\frac{I}{I_0} = e^{-aCL} , \qquad (6)$$

where I/I_0 is the transmittance through the path length L (m). The mass extinc-

tion coefficient a (m^2/g) can be considered the removal (scattering and absorption) cross section per unit mass and must be determined empirically. It is a function of aerosol particle size distribution, particle shape, and electromagnetic wavelength. 13 The product αC (m-l) can be considered as the volumetric extinction coefficient, σ, and is the value that is generated by the LCM support software.
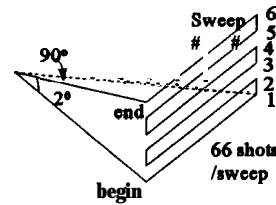

Fig. 2. Schematic of the sweep pattern of the DREV Laser Cloud Mapper.

A weakness inherent in the LCM is that aerosol clouds may be sufficiently dense to provide near-complete extinction of the laser pulses. In such cases, only the near side of the cloud (closest to the LCM) will provide reflections which can be converted to extinction coefficients or concentrations. For this reason, more confidence is placed in measurements from sparse clouds than from ones which are optically dense.

Data Handling

The backscatter signal strengths collected by the LCM are converted by support software to volume extinction coefficient values. With the vast amount of data available per six-sweep scan, locating extinction values for the aerosol cloud was not straightforward. A thresh-

old value of 2.00×10^4 was used as a filter, and the **backscatter** resulting from reflections from the natural background, such as trees and terrain features, had to be removed. The coordinate system was defined by having the x axis pass through the cloud **centroid,** with the origin of the **cartesian** coordinate system determined as the cloud **centroid** for the first scan.

In **addition,** the data had to be arranged into a format suitable for training and testing ANNs and in validating both the ANN and Gaussian models. The **validation** set was selected as the last scan of one of the tests and was used only for blind **validation,** i.e., it was not used for training or testing. This set contained 1103 data points.

The extinction coefficients available for analysis ranged over four orders of magnitude (10^{-1} to 10^{-4}). Consequently, a logarithmic transformation was performed. Another adjustment was to balance the distribution of the data points over the range of extinction coefficients. This was achieved by randomly choosing points from the residual data set (less the **validation** set) to fill a number of evenly wide data sub-range brackets throughout the available extinction coefficient range. The purpose of this was to preclude an inadvertent biasing of the ANN towards any particular sub-range of values.

ANN Development

The choice of ANN **paradigm,** back propagation, was based on the wide use and general success it has exhibited in solving prediction problems. [14,15] The version used was NeuralWare's Neural-Works Professional **II/Plus**[16] with the edbd or extended delta-bar-delta **option,** which included dynamic adjustments to momentum and learning coefficients.

Another parameter investigated during ANN training was in choosing the input variables and parameters. The choice was limited somewhat by the nature of the data available, i.e., only two tests were used, with both conducted under relatively **uniform** meteorological conditions. Consequently, time of day and temperature w-ere d-eemed insignif-cant. Further, specific details of the grenades were not available, so obvious inputs to describe the aerosols produced, such as total weight of aerosol dispersed, the burn/source emission rate and particle size **distribution,** were not included. Also, such values as ambient temperature, barometric pressure, relative humidity and thus chemical kinetics data were not available and so were not considered. The variables actually considered are contained in Table 1.

Table 1. Input variables for the artificial neural network, with relative significance determined by sensitivity analysis and normalized to 'most significant.

Input Variable or Parameter	Relative Significance
Downwind distance (x)	1
Crosswind distance (y)	.78
Wind velocity	.35
Cloud deviation from centreline	.28
Type of grenade	.24
Time from launch	.23
Type of grenade	.22
Vertical distance (z)	.16
Pasquill category	.05

The final area of ANN development was the architecture of the network. Here the choice was in the number of hidden processing elements or nodes and their arrangement into one or two hidden layers. The optimal architecture

arrived at, after significant investigation, was 9-9-4-1, or nine input nodes, nine nodes in the first hidden layer, four in the second hidden layer and one node in the output layer (logarithm of the volumetric extinction coefficient). In general, the fewer the number of hidden nodes the better, as this forces the model to generalize and not memorize. The larger number used in this application may well be due to synergistic effects among some of the inputs.

Model Performance

The metric chosen to assess model performance was the average relative error, E, of the validation set, which was applied to both the Gaussian and the ANN models. This is determined by

$$E = \sum_n \left| \frac{\sigma_m - \sigma_p}{\sigma_m} \right|, \qquad (7)$$

where $\sigma_{m,i}$ is the LCM-measured volumetric extinction coefficient (m-l) and $\sigma_{p,i}$ is the model-predicted volumetric extinction coefficient (m-l).

A scatter plot of the Gaussian plume model can be seen in Fig. 3. Perfect agreement between measurement and prediction would be depicted by points falling along a diagonal with a slope of 1. It can be seen that most of the data are overpredicted by the Gaussian model while a significant number (the points aligned at the bottom) were underpredicted. Here these data were arbitrarily assigned predicted values of -4.7, so that they would be shown graphically without greatly distorting the depiction of the majority of the data points. In fact, the points, in some cases, were several

orders of magnitude below the measured values.

The relative error for the Gaussian model was determined to be 14.5, or almost 1.5 orders of magnitude, due largely to the underpredictions at the cloud extremities (Table 2). It is worth noting that the Gaussian plume model is the combination of 12 separate plume models, one for each grenade launched. It should also be noted that the puff model was also investigated, but the predictions were not as good as those from the plume model, probably because the standard deviations are better characterized for the more common plume model.
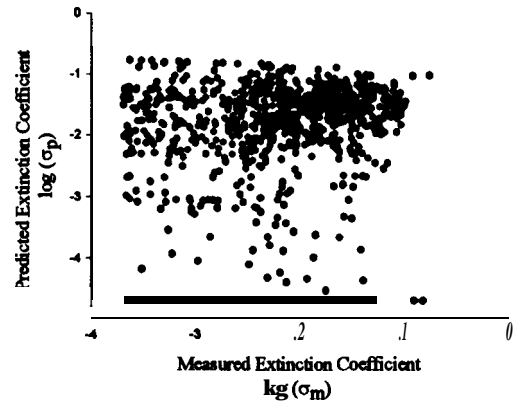

Fig. 3. Scatter plot of Gaussian plume model predictions of logarithms of volumetric extinction coefficients vs logarithm of LCM measured values for validation set of 1103 values.

Table 2. Comparison of average relative errors of ANN and Gaussian plume predictions for the validation set.

	Neural Network	Gaussian Plume
Error	1.78	14.5

The **performance** of the ANN model can be seen in Fig. 4. The data are much better distributed about the ideal line of slope 1. The relative error of 1.78 for the validation set is quite reasonable, indicating that actual predicted values are, on average, within a factor of 2 of the measured values Table 2. This difference, of course, is much smaller (0.25) when the logarithmic values are considered, as depicted in Fig. 4.
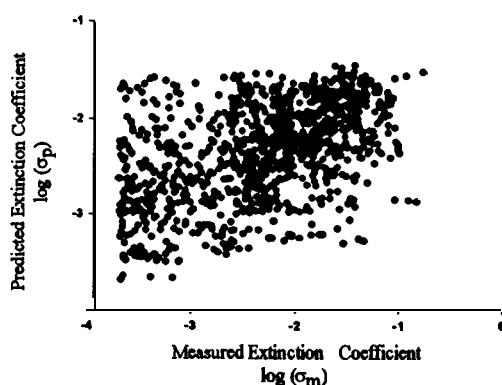

Fig. **4.** Scatter plot of artificial neural network prediction of logarithm of volumetric extinction coefficient vs logarithm of LCM measured values for validation set.

Noteworthy also is the fact that the ANN model does not predict extinction coefficients above about 0.04 m^{-1}, even though some measured values were as high as 0.7 m^{-1}. This is probably due to the generalizing nature of the ANN, which was trained on a paucity of data above 0.1 m^{-1}.

It should also be borne in mind that the training, test and validation sets were limited to two tests (i.e., two separate volleys of 12 grenades each). Although more data were available, there was insufficient **information** on grenade compositions and meteorological condi-

tions. Consequently only the two firings were chosen for **evaluation,** and data from most of the tests conducted, which could not be sufficiently characterized, were not used.

The ANN model developed, **then,** is not a generalized **predictor** of aerosol dispersion. Rather, it is a proof of principle that an artificial neural **network,** with the inherent flexibility of choice of input variables can provide a good prediction of a fairly complex and chaotic **phenomenon,** the dispersion of aerosols in the atmosphere.

Examination of the connecting weight distribution of the ANN and conducting a sensitivity analysis yields the relative significance of the input variables, as shown in Table 1. Not surprisingly, especially in light of the similarity of the atmosphere conditions, the most significant variables are spatial ones in the ground plane. At the other extreme, the **Pasquill** stability **category,** which for one test was A-B and for the other was B-C, had predictably little influence. Similarly, the vertical distance above ground had little influence, again probably due to the fact that a height of only 14m was covered by the LCM. The other variables, the type switches which indicated which grenade was used, the time after launch and the degree of cloud meander (deviation of cloud axis **from** wind direction at cloud **centroid) all** had, as could be anticipated, some influence. Also, as might be expected, the wind velocity was an important influence.

Fig. 5 depicts the extinction coefficient or concentration profile along a line of sight 1.5 m above ground at 110 m from the source and normal to the cloud longitudinal **axis/wind** direction. It can be seen that, in this instance, the actual cloud had a **bimodal distribution,**

while the Gaussian model, a composite of 12 separate Gaussian distributions, depicts a smooth monomodal distribution. (Since the source strength was unknown, the Gaussian was normalized to the highest measured extinction coefficient). The ANN model, however, did reproduce the bimodal distribution.
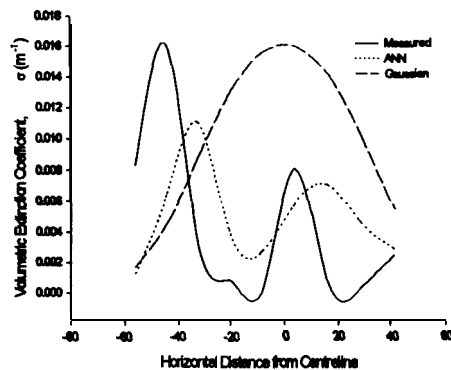

Fig. 5. Distribution of predicted and measured extinction coefficients along a line of sight 110 m from the source and normal to the cloud, 1.5 m above ground for a composite RP/metal flake cloud from a 12 grenade volley.

Conclusion

From the work reported, a number of conclusions can be drawn. The data available from the DREV trials are incomplete. The grenade compositions and values of a number of meteorological parameters were not available from the trial reports. This limited both the ability to compare different grenades and to provide a full input space to train neural networks.

On the limited data set examined, the ANN model provided better predictions than the Gaussian plume model (which, in turn, was better than the

Gaussian puff model). Finally, the LCM has been found to provide time-sensitive spatial aerosol concentration distributions within clouds.

Future work being planned includes trials involving single grenades (both L8 type RP and inert simulations of M76 type metal flake) to ensure that the LCM laser signal is not extinguished by the cloud, i.e., that signals are received from throughout the cloud, and to permit comparison with COMBIC (Combined Obscuration Model for Battlefield Induced Contaminants),[17] a Gaussian-based code developed by the U.S. Army Research Laboratory.

Acknowledgement

References

1. Evans, A. J., "A Preliminary Assessment of the DREV Laser Cloud Mapper," CDE Porton Down, UK (1982).
2. Evans, B. T.N., Kluchert, R.E. and Levesque, R. J., "Field Evaluation of a Canadian Laser Cloud Mapper and Canadian IR Screening Aerosols," DREV 4271/82, Valcartier, Quebec (1983).
3. Evans, B. T.N., Roy, G. and Ho, J., "The Detection and Mapping of Biological Simulants: Preliminary Biological Results," DREV 4480/89, Valcartier, Quebec (1989).
4. Twardawa, P., Roy, G., Evans, B.T.N. and Vallee, G. "Cold Weather Evaluation of Visual and Infrared Screening Grenades for the Self-Protection of Armoured Fighting Vehicles," DREV R-4632/91, Quebec (1991).
5. Csanady, G. T., *Turbulent Diffusion in the Environment,* D. Reidel Publishing Co., Dordrecht Holland. (1973).
6. Pasquill, F and Smith, F. B., *Atmospheric Diffusion, 3rd Ed.,* Ellis Horwood Ltd., ChichesterUK(1983).
7. Turner, D.B., *Workbook of Atmospheric Dispersion Estimates, 2nd Ed.,* Lewis, Boca Raton (1994).
8. Hanna, S. R., Briggs, G.A. and Hosker, R.P., "Handbook on Atmospheric Diffusion," DOE/TIC-l 1223, Springfield VA (1982).
9. Rumelhart, D.E. and McClelland, J. L., *Parallel Distributed Processing - Explorations in the Microstructure of Cognition Vol 1: Foundations,* MIT Press, Cambridge MA (1986).
10. Roy, G., Evans, B.T.N. and Gilbert, J., "h Overview of the Laser Cloud Mapper Results from JUSCAN-OASIS (Smoke Week IX) Trials," DREV R-4676/92, Quebec (1992).
11. Roy, G., Bonnier, D., De Villers, Y., Couture, G., Hutt, D. and Vallee, G., "Canadian National Report on the SOCMET Winter Test Held At DREV, Canada in March 1993," DREV-TM-9408, Quebec (1994).
12. Roy, G., Couture, G. and Vallee, G., "Canadian National Report on the SOCMET Summer Test Held at E. T. B.S., France in September 1993," DREV-TM-9434, Quebec (1995).
13. Evans, B. T.N., "An Interactive Program for Estimating Extinction and Scattering Properties of Most Particulate Clouds," MRL-R-1 123, Materials Research Laboratory, Ascot Vale, Victoria, Australia (1988).
14. Patterson, D. W., *Artificial Neural Networks,* Prentice Hall, Singapore (1996).
15. Baughman D.R. and Liu, Y.A., *Neural Networks in Bioprocessing and Chemical Engineering,* Academic Press, San Diego (1996).
16. NeuralWare, Inc., *NeuralWorks Professional 11/Plus,* Version 5.0. Computer Software. (1991).
17. Ayres, S.D. and DeSutter, S., *Combined Obscuration Model for Battlefield Induced Contaminants (COMBIC92) Model Documentation,* U.S. Army Research Laboratory, White Sands Missile Range NM (1995).